Probability Comprehensive Exam
Spring 2018

Student Number: ☐

Instructions: Complete 5 of the 9 problems, and **circle** their numbers below – the uncircled problems will **not** be graded.

1 2 3 4 5 6 7 8 9

Write **only on the front side** of the solution pages. A **complete solution** of a problem is preferable to partial progress on several problems.

1. Let $\{X_n\}$ be a sequence of independent identically distributed random variables with exponential distribution (in other words, $X_n \geq 0$ a.s. and $\mathbb{P}\{X_n \geq t\} = e^{-\lambda t}, t \geq 0$ for some $\lambda > 0$). Prove that
$$\limsup_{n\to\infty} \frac{X_n}{\log n} < \infty \text{ a.s.}$$

Solution: For $C > 0$, $\mathbb{P}\{X_n \geq C \log n\} = e^{-C\lambda \log n} = n^{-C\lambda}$. For $C > \lambda^{-1}$,
$$\sum_{n\geq 1} \mathbb{P}\{X_n \geq C \log n\} = \sum_{n\geq 1} n^{-C\lambda} < \infty.$$
By Borel-Cantelli Lemma,
$$\mathbb{P}\{X_n \geq C \log n \text{ infinitely often}\} = 0,$$
implying the claim.

2. Suppose f is a continuous function on $[0, 1]$. Use the Law of Large Numbers to prove that
$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 f((x_1 \dots x_n)^{1/n}) dx_1 \dots dx_n = f\left(\frac{1}{e}\right).$$

Solution: Let $X_1, \dots, X_n$ be i.i.d. random variables with uniform distribution in $[0, 1]$. Then
$$\int_0^1 \cdots \int_0^1 f((x_1 \dots x_n)^{1/n}) dx_1 \dots dx_n = \mathbb{E} f((X_1 \dots X_n)^{1/n})$$
$$= \mathbb{E} f\left(\exp\left\{\frac{\log X_1 + \cdots + \log X_n}{n}\right\}\right).$$
By the Strong Law of Large Numbers,
$$\frac{\log X_1 + \cdots + \log X_n}{n} \to \mathbb{E} \log X_1 = \int_0^1 \log x dx = -1 \text{ as } n \to \infty \text{ a.s.}$$
By continuity of f and Lebesgue dominated convergence theorem,
$$\mathbb{E} f\left(\exp\left\{\frac{\log X_1 + \cdots + \log X_n}{n}\right\}\right) \to f(\exp\{-1\}),$$
implying the result.

3. Let X, Y be random variables with $\mathbb{E}|X| < \infty, \mathbb{E}|Y| < \infty$. If $\mathbb{E}(X|Y) = Y$ and $\mathbb{E}(Y|X) = X$ a.s., then $X = Y$ a.s. Prove it.

Solution: Let f be a uniformly bounded strictly increasing function. It follows from the assumptions that

$$\mathbb{E}(X-Y)f(Y) = \mathbb{E}\mathbb{E}(X-Y|Y)f(Y) = \mathbb{E}(\mathbb{E}(X|Y)-Y)f(Y) = \mathbb{E}(Y-Y)f(Y) = 0.$$

Similarly, $\mathbb{E}(X-Y)f(X) = 0$, which implies

$$\mathbb{E}(X-Y)(f(X)-f(Y)) = 0.$$

Since f is strictly increasing, $(X-Y)(f(X)-f(Y)) \geq 0$ and, moreover, $(X-Y)(f(X)-f(Y)) = 0$ if and only if $X = Y$. Therefore, we have $(X-Y)(f(X)-f(Y)) = 0$ a.s., implying $X = Y$ a.s.

4. Let $X_1, \ldots, X_n$ be i.i.d. random variables with mean μ and variance $\sigma^2 < +\infty$. Let f be a function continuously differentiable at the point μ. Prove that the sequence of random variables

$$n^{1/2}\left(f\left(\frac{X_1 + \cdots + X_n}{n}\right) - f(\mu)\right)$$

converges in distribution to a normal random variable. What is the mean and the variance of the limit?

Solution: Let $Y_n = n^{1/2}(\frac{X_1+\cdots+X_n}{n} - \mu)$. By the Central Limit Theorem, Y_n converges in distribution to a normal random variable Y with mean zero and variance σ^2 as $n \to \infty$ and, by the Law of Large Numbers, $n^{-1/2}Y_n \to 0$ as $n \to \infty$ in probability. By the first order Taylor expansion,

$$f(x) = f(\mu) + f'(\mu)(x-\mu) + r(\mu; x-\mu)(x-\mu),$$

where $r(\mu;\delta) \to 0$ as $\delta \to 0$. Therefore,

$$\begin{aligned} n^{1/2}\left(f\left(\frac{X_1 + \cdots + X_n}{n}\right) - f(\mu)\right) &= n^{1/2}(f(\mu + n^{-1/2}Y_n) - f(\mu)) \\ &= f'(\mu)Y_n + r(\mu; n^{-1/2}Y_n)Y_n. \end{aligned}$$

Since $f'(\mu)Y_n$ converges in distribution to $f'(\mu)Y$ and $r(\mu; n^{-1/2}Y_n)Y_n$ converges in probability to 0, we can conclude that $n^{1/2}\left(f\left(\frac{X_1+\cdots+X_n}{n}\right) - f(\mu)\right)$ converges in distribution to a normal random variable with mean 0 and variance $(f'(\mu))^2\sigma^2$.

5. Let $X_1, \ldots, X_n, \ldots$ be i.i.d. random variables with $\mathbb{E}X_1 = 0$ and $\mathrm{Var}(X_1) = 1$. Let $S_n = X_1 + \cdots + X_n$. Prove that

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} = +\infty.$$

Solution: By the Central Limit Theorem, for all $A > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left\{ \frac{S_N}{\sqrt{N}} \geq A \right\} = \mathbb{P}\{Z \geq A\} > 0,$$

where Z is a standard normal random variable. Therefore,

$$\mathbb{P}\left\{ \limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} \geq A \right\} = \mathbb{P}\left\{ \lim_{N\to\infty} \sup_{n\geq N} \frac{S_n}{\sqrt{n}} \geq A \right\}$$

$$= \lim_{N\to\infty} \mathbb{P}\left\{ \sup_{n\geq N} \frac{S_n}{\sqrt{n}} \geq A \right\} \geq \lim_{N\to\infty} \mathbb{P}\left\{ \frac{S_N}{\sqrt{N}} \geq A \right\} > 0.$$

On the other hand, for all $m \geq 1$,

$$E = \left\{ \limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} \geq A \right\} = \left\{ \limsup_{n\to\infty} \frac{X_m + \cdots + X_n}{\sqrt{n}} \geq A \right\} \in \mathcal{F}_m = \sigma(X_m, X_{m+1}, \ldots).$$

Thus, by Kolmogorov's Zero-One Law, $\mathbb{P}(E)$ is either 0, or 1. Since $\mathbb{P}(E) > 0$, we have

$$\mathbb{P}\left\{ \limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} \geq A \right\} = 1$$

for all $A > 0$, implying the claim.

6. Let (X_n) be an i.i.d. sequence of random variables with

$$\mathbb{P}(X_n = 1) = 1/2 = \mathbb{P}(X_n = -1).$$

Let (Y_n) be a bounded sequence of random variables such that $\mathbb{P}(Y_n \neq X_n) \leq e^{-n}$. Show that

$$\frac{1}{n}\mathbb{E}(Y_1 + \cdots + Y_n)^2 \to 1 \text{ as } n \to \infty.$$

Solution: For $n \geq 0$, let A_n be the event

$$A_n = \{X_k \neq Y_k \text{ for some } k \geq n^{1/4}\}.$$

Then

$$\mathbb{P}(A_n) \le \sum_{k \ge n^{1/4}} \mathbb{P}(X_k \ne Y_k) \le \sum_{k \ge n^{1/4}} e^{-k} \le C_1 e^{-n^{1/4}}.$$

Now write $S_n = X_1 + \cdots + X_n$ and $T_n = Y_1 + \cdots + Y_n$. Using the Cauchy-Schwarz inequality,

$$\begin{aligned}
\mathbb{E}\left(\frac{S_n}{\sqrt{n}} - \frac{T_n}{\sqrt{n}}\right)^2 &= \frac{1}{n}\mathbb{E}(S_n - T_n)^2 \mathbf{1}_{A_n} + \frac{1}{n}\mathbb{E}(S_n - T_n)^2 \mathbf{1}_{A_n^c} \\
&\le \frac{1}{n}\sqrt{\mathbb{E}(S_n - T_n)^4 \mathbb{P}(A_n)} + \frac{1}{n}\mathbb{E}(S_n - T_n)^2 \mathbf{1}_{A_n^c} \\
&\le C_2 \frac{1}{n} n^2 e^{-(1/2)n^{1/4}} + \frac{1}{n}\mathbb{E}(S_n - T_n)^2 \mathbf{1}_{A_n^c}.
\end{aligned}$$

On the event A_n^c, one has $|S_n - T_n| \le C_3 n^{1/4}$, so the second term above is bounded by $\frac{1}{n} C_3^2 \sqrt{n}$. We obtain an overall bound of

$$C_2 n e^{-(1/2)n^{1/4}} + C_3^2/\sqrt{n},$$

which goes to 0 as $n \to \infty$. Therefore $S_n/\sqrt{n} - T_n/\sqrt{n} \to 0$ in L^2. Because $\|S_n/\sqrt{n}\|_2 = 1$ for all n, the triangle inequality gives

$$|\|T_n/\sqrt{n}\|_2 - 1| \le \frac{1}{\sqrt{n}}\|T_n - S_n\|_2 = \sqrt{\frac{1}{n}\mathbb{E}(S_n - T_n)^2} \to 0.$$

In other words, $\frac{1}{n}\mathbb{E}T_n^2 \to 1$.

7. Let F_n, F be distribution functions such that $F_n \to F$ weakly. If F is continuous, show that

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \to 0.$$

Solution: Let $\epsilon > 0$. Because F is continuous, we may choose a finite collection of points $x_1, \ldots, x_K$ such that $F(x_i) = i\epsilon/3$. (Here K is chosen as $\lfloor 3/\epsilon \rfloor$.) Because $F_n \to F$ weakly, $F_n(x) \to F(x)$ at each continuity point x of F, and since F is continuous, $F_n(x) \to F(x)$ for all x. Thus we may choose N such that $n \ge N$ implies that $|F_n(x_i) - F(x_i)| < \epsilon/3$ for all $i = 1, \ldots, K$.

Now if $n \ge N$ and x is such that $x \in [x_i, x_{i+1}]$, one has

$$F(x_i) - \epsilon/3 < F_n(x_i) \le F_n(x) \le F_n(x_{i+1}) < F(x_{i+1}) + \epsilon/3,$$

and

$$F(x_i) \le F(x) \le F(x_{i+1}).$$

This means that both $F(x)$ and $F_n(x)$ are in the interval $(F(x_i) - \epsilon/3, F(x_{i+1}) + \epsilon/3)$, and so

$$|F_n(x) - F(x)| < F(x_{i+1}) - F(x_i) + 2\epsilon/3 = \epsilon.$$

On the other hand, if $x < x_1$, $0 \le F_n(x) \le F_n(x_1) < F(x_1) + \epsilon/3$ and $0 \le F(x) \le F(x_1)$, giving $|F_n(x) - F(x)| < 2\epsilon/3$. Similarly, if $x > x_K$, then $F(x_K) - \epsilon/3 < F_n(x_K) \le F_n(x) \le 1$ and $F(x_K) \le F(x) \le 1$, giving $|F_n(x) - F(x)| < 2\epsilon/3$. Putting the three cases together, for $n \ge N$, $\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| < \epsilon$.

8. Let (X_n) be an i.i.d. sequence of random variables. Show that $\mathbb{E}(X_1)^2 < \infty$ if and only if for every $c > 0$, $\mathbb{P}(|X_n| \ge c\sqrt{n}$ infinitely often$) = 0$.

Solution: Suppose first that $\mathbb{E}(X_1)^2 < \infty$. Then for any $c > 0$,

$$\sum_{n\ge 0} \mathbb{P}(|X_n| \ge c\sqrt{n}) = \sum_{n\ge 0} \mathbb{P}(X_n^2 \ge c^2 n) = \sum_{n\ge 0} \mathbb{P}(X_1^2/c^2 \ge n).$$

We can write the right side using monotone convergence as

$$\mathbb{E}\sum_{n\ge 0} \mathbf{1}_{\{X_1^2/c^2 \ge n\}} = \mathbb{E}\left(\left\lfloor \frac{X_1^2}{c^2} \right\rfloor + 1\right) \le 1 + \frac{1}{c^2}\mathbb{E}X_1^2 < \infty.$$

So by the Borel-Cantelli lemma, $\mathbb{P}(|X_n| \ge c\sqrt{n}$ infinitely often$) = 0$.

Conversely, if $\mathbb{P}(|X_n| \ge c\sqrt{n}$ infinitely often$) = 0$, since the variables (X_n) are independent, these events are also independent, and so the Borel-Cantelli lemma (and reversing the above computation) gives

$$\infty > \sum_{n\ge 0} \mathbb{P}(|X_n| \ge c\sqrt{n}) = \mathbb{E}\left(\left\lfloor \frac{X_1^2}{c^2} \right\rfloor + 1\right) \ge \frac{1}{2}\mathbb{E}(X_1^2/c^2 + 1).$$

This implies that $\mathbb{E}X_1^2 < \infty$.

9. Find an example of a random variable X with a density function but whose characteristic function ϕ_X satisfies

$$\int_{-\infty}^{\infty} |\phi_X(t)| \, \mathrm{d}t = \infty.$$

Solution: Let X be exponential with mean 1. Then its characteristic function is

$$\phi_X(t) = \mathbb{E}e^{itX} = \int_0^\infty e^{itx}e^{-x}\ \mathrm{d}x = \frac{1}{1-it} = \frac{1+it}{1+t^2}$$

Therefore

$$|\phi_X(t)| = \frac{1}{1+t^2}\sqrt{1+t^2} = \frac{1}{\sqrt{1+t^2}} \geq \frac{1}{|t|},$$

and so

$$\int_{-\infty}^\infty |\phi_X(t)|\ \mathrm{d}t \geq \int_1^\infty \frac{\mathrm{d}t}{t} = \infty.$$